Diversification of Regional Dialect

Maureen Ballard,[1] Robert Stojnić,[2] and William T. Gibson[3]

[1] *Instituto Balseiro, S.C. de Bariloche, Argentina*
[2] *Faculty of Technical Sciences, Novi Sad, Serbia and Montenegro*
[3] *Harvard University, Boston, Massachusetts, USA*

The study of language evolution has been addressed at many different levels from early human acquisition to childhood learning. Here we look at a further refined scale and examine regional accent acquisition and diversification through small pronunciation perturbations among speakers who have already learned a single established lexicon. The model is built primarily on the constraints that speakers interact on a network of social connections through a vocabulary divided into word groups. The manner in which these word groups interact is through imitation of a phase variable which is extracted by mapping the anatomical vowel sound chart onto the unit circle. We then study the implications of linguistically motivated coupling matrices between speakers. Some coupling matrices considered here indeed produce diversity of pronunciation within each word group. Another source of regional accent diversity arises when speakers are considered to have a nonzero reception/imitation error. While many social pressures and particular interaction functions are for now neglected in the model, we discuss and explore features that linguistics perceive to govern language evolution on the scale of accent diffusion, and directions of further study.

I. INTRODUCTION

The original crumb of spoken word was lost long ago. It fled the scene, it spontaneously multiplied, it was immediately eaten by it's utterer. Whatever the fate of this little lost phoneme, we are left today with the questions concerning its legacy. One such example being, how do speakers affect and alter each other's speech through linguistic interaction? Among the multitudes of other related questions that might intrigue a person, we here attempt to address an mere aspect of this one. Typically language is learned when we are young, with the lilt and style of those who influence us most. But this does not imply that we are locked for life into that manner of speaking. Living in different regions and interacting with different people, the language we speak is evolving as tumbles from our lips. Is accent and speech style something unique and permanent to each of us, or as social creatures, do we assimilate and imitate our speech environment? In this work, we began with a single question, which immediately branched into many, addressing the various social influences and pressures that affect each linguistic encounter. With such considerations, we now ask, under certain restraints, how might diversification in regional accent emerge from a single, standardized mother language? What are some of the crucial mechanisms through which we affect and alter each other's speech through linguistic interaction? What social pressures are important to linguistic diffusion and propagation? How stable is a given regional accent configuration under small accent perturbations?

In constructing the following model, we have tried to isolate the most important social and linguistic pressures as expressed by those knowledgeable of real linguistic systems. We begin by asserting that there are speakers which interact with a limited group of other speakers. That is, the population is not perfectly mixed, but that in a large enough population, members only speak with a subset of the total population. For this reason, interactions occur on the substrate of a social network. As a second basic principal of the model, we propose a dynamic variable to characterize the means by which pronunciation changes occur. This pronunciation variable tracks the vowel sound pronunciation of all speakers and for every word group. It thus may be used to measure the speech patterns of the entire population. Among the many social pressures that dictate how people interact with each other, we begin with one of the most simple aims of communication. We assume that speakers consciously or otherwise, attempt to optimize transmission and reception. This dichotomy illustrates the speaker's need to assimilate or synchronize forms of speech with other speakers, but also the necessity for distinction and clear expression. Based upon the network substrate, pronunciation phase variable, and these two 'opposing forces', we have built into the model a few other social pressures that strike us as highly pertinent. Features are added, one by one to the idealized system, as we observe how each affects the overall dynamics, especially in the equilibrium state. The model may continue to be constructed in this manner by future curiosity seekers (though correspondence with reality is never guaranteed).

With some numerical analysis, we confirm a few trivial cases of pronunciation assimilation. However, with certain linguistically motivated changes in the parameters governing the couplings between speakers, we find a diversification in the pronunciation of vowel sounds between the speakers.

II. HUMAN COMMUNICATION

Consider the quantity of monosyllabic sounds that a human being is capable of producing. Each of these sounds is known in linguistics as a phoneme, which together represent the most basic units of sound used to compose words and morphemes [1]. As an exercise in combinatorics, now consider all possible combinations of all such phonemes. This quantity of words, phonetic combinations, that human utterance is physiologically capable, is immense, even when word length is assumed to be finite. Fortunately or otherwise, languages compose small subsets of this larger set of all possible words. Thought of in this manner, given the wide variety of sounds that each human might choose to produce, how do groups of humans agree upon and use sometimes to exclusion, one of these smaller word sets? A lexicon, loosely defined as a dictionary or mental vocabulary, allows people to map utterances to meanings in a consistent manner. Meaning-utterance maps have primarily been studied as an abstraction of the learning process that children experience as they explore and define their environments, or as a means of language acquisition in early humankind [3, 4]. The idea begins with each speaker presenting an utterance in relation to a given concept (object, verb, etc.). Whether or not communication is the explicit goal, speakers tend to imitate each other as they interact and rename or re-utter, each concept or meaning. If interaction rates are strong enough, meaning-utterance maps converge and a definition is built. Finally, an entire lexicon is agreed upon for a group of people. This theory suggests a phenomenon unique to *interacting* individuals, known as self-organization. Components interact and adapt until a sort of synchrony is achieved. While synchronization is a common way to capture more primitive forms of communication, such as light pulsation in Nile Valley fireflies [5], perhaps it may be applied similarly to human lexicons.

Language is a highly complex process, serving as the most distinctive aspect of the human species [6, 7]. While other animals are capable of communication, no other has developed such a physiologically complete and grammatically consistent means of transmission, reception, and resultant social organization, as have hominids [6, 8]. Each lexicon, of the many that have walked the earth, is particular in some manner. Some differ in origin, and thus share few to zero similarities in grammar, structure, or sound. Whereas other lexicons may have evolved from the same root language, and thus exhibiting many common traits. Further, a single lexicon might have many distinct forms of pronunciation of the same words. These various lexical comparisons could be thought of as different stages in language evolution, one ultimately leading to the next.

In this study we will focus on slight differences in pronunciation or accent, of a given language. This way, all interactions occur through an already established meaning-utterance map, and speakers do not actually learn new words. By restricting communication to a single lexicon, we study how certain vowel sounds influence each other. Without the additional noise of other languages (although these perturbations would compose a more realistic situation), one may consider the lexicon to have a constant size, and only be subject to small alterations of the standardized pronunciation. In grammatical surveys done by Labov, slight differences in vowel phonemes are found to be spatially correlated, see Fig. 1. These differences suggest several interesting aspects of human communication that have motivated many theories. First, that a language of precise origin and pronunciation may diffuse and evolve over space and through time. Second, that social interaction structures keep these variations regionally intact within a certain timescale. These observations have guided several questions addressed in this work as well. In the context of interacting vowel sounds, we will examine the diffusion and propagation of regional accent. Further, we will look at the effects of the underlying sociological topologies woven by the individuals who speak with each other, and thus influence each other. Much work has been done with respect to linguistic diffusion [6–9]. Thus, before we delve into the proposed model, some background in language propagation and evolution is necessary.

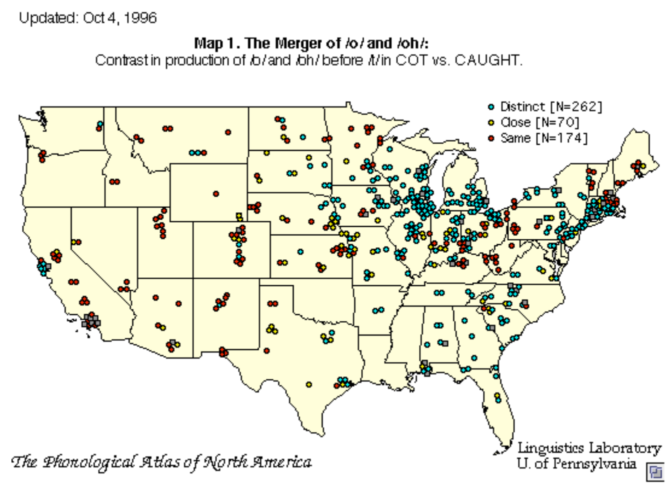

FIG. 1: The graph show the geographical distribution of mergers of /o/ and /oh/ as presented by Labov. The mergers show geographical confinement.

III. LINGUISTIC DIFFUSION

Diffusion is a concept rudimentary to many scientific disciplines. From a statistical analysis, the process of diffusion is the result of many random steps taken by agents as they wander about and interact momentarily with other elements in the environment. This spooky concept, first witnessed in the eyepiece of botanist Robert Browns microscope was clarified at the same time, but on different sides of the globe, by Albert Einstein and William Sutherland. Since then, this phenomenon and following conclusions have been witnessed repeatedly in nature. However, what is observed in nature tends to be on a macroscopic scale and thus the individual, microscopic movements are visually averaged over. Populations of molecules, chemicals, cells, animals, etc. are observed to diffuse, interact, deplete, reproduce, and move, on this larger scale as whole. Normalized population densities thus characterize the behavior of a statistical ensemble of individually acting components, through mean values, variances, and other moments. In linguistics, the process of diffusion is used to understand the propagation and subsequent evolution of language over a population of speakers.

Given the complexity of social interactions, language transmission is likely to be correlated to many causes and situational influences. Susceptibility to language acquisition is clearly connected with age, but to what extent? Different social pressures are exerted on different people at different moments in their lives. Personality, gender, economic status, demographics, education, and many other factors influence our speech profiles. To address this complication, linguists have studied human language statistically and physiologically, with the hope of extracting the most important features of linguistic diffusion. One large distinction is between language that is acquired early in life from parents and other members of the previous generation, and language acquired through interactions with friends and peers throughout the lifetime. These are referred to as vertical and horizontal transmission, respectively [10, 11]. The medium of transmission is speech, what is actually transmitted are linguistic innovations which include: slang, pronunciation, new words, and word frequency among others contagious modes of expression.

Depending on what one hopes to extract from the linguistic model, certain social pressures ought to be considered in the diffusive process. Linguistic *style* and *code switching* illustrate the use of different slang or language, each within its appropriate setting [8, 12, 13]. By consciously or unconsciously doing this, people recognize that some topics or situations require a more formal presentation, whereas others ought to be more casual, intelligent, sarcastic, competitive, consoling, etc. The implications of *social status* are often woven into the spoken language as well. This linguistic pressure not only acts in a locally diffusive manner among those participating in the same class activities, but also induces a diffusive bias by those who wish to practice expressions of another class. Parents tend to speak to children with an accent pertaining to higher class in attempt to increase the child's educational prospects and thereby inducing a biased vertical diffusion of pronunciations. It should be noted that although social status is often expressed through language, there are many pronunciation differences that are not associated directly with class [8]. *Gender* and *age* also affect the conscience and unconscious choices of pronunciation and vocabulary usage, however, this depends heavily on culture. Interestingly, children seem to be the ones that assimilate linguistic change. That is, adults vertically transmit inconsistent innovations to which children apply the rules, implicit to a given innovation, to other words in the lexicon to make a consistent vocabulary. The youth also tend to adopt more slang and nonstandard forms of speech, thus introducing many changes in accent and vocabulary into the lexicon, however transient they may be [8, 12].

Finally, *social networks* are considered to have the most profound influence on the speech of a given individual. Not only the social topology, but also the nature of the interactions governed by the social pressures listed above: gender, age, social status, etc., are encoded in the network. It is thus postulated that the interaction network might provide the most concentrated means to study propagation of linguistic innovation. Groups of people who interact are frequently inhomogeneous and come from different backgrounds (i.e. Puerto Ricans and African Americans in NYC), each with their own accent particularities. These interactions then lead to deviations from their background pronunciation scheme, and to mixtures thereof. In NYC, it has been observed that Puerto Ricans who were in contact with African American speakers deleted final -t/d more often (e.g. fas' computer, gol' sunset) than other Puerto Ricans. A second study was done on local dialect by examining the correlations between speech and familial connections of two females in Belfast. The local pronunciation and slang were much more prevalent in the woman who had many local kin. Whereas the other woman who did not have many local relatives, spoke a relatively standard English. In Brazil, it was shown that migrants from rural areas have different assimilation patterns with respect to gender. In contrast to other studies done on gender correlation to speech innovation, females tended to assimilate to urban speech more slowly than men. This was interpreted by the different social networks of males, who have work ties in the city, and females (who migrate with them), who have fewer urban social contacts [8]. In this work, we will study the effects of different social weights between pairs of speakers (coupling strengths) within the social network as an important part of the final analysis. Though these coupling strengths are initially presented in a general manner, they may later be interpreted and tailored to some of the social pressures that have been discussed here.

A. The Macro Models

As innovations are exchanged on the microscopic level (between pairs of speakers), the global characterization of a given language group is fluctuating as speakers explore and decline new mannerisms and accents. Some innovations remain, or at least remain locally, thus creating some distinction between groups on a global level. In the *wave theory* as described in [7], changes are thought to nucleate in one region and then spread in concentric, circular rings known as *isoglosses*. With this, one may envision the growth and collision of many rings as they spread from their respective sources of linguistic change. The surface of a pond during a rainstorm is a helpful analogy. This set of dynamic and interacting isoglosses, arising from independent events, is typically subject to environmental and demographic variation in the landscape [4, 7, 11].

A famous example of such a dialect continua is the Rhenish fan. The two varietes of German, the Highand Low-German are separated by the so-called Benrath-Urdingen isogloss running east-west.High-German exhibits the sound shift from the original Low-German, $\{p, t, k\} \longrightarrow \{(p)f, s, x\}$. For instance, in Low-German one says *dorp*, while in High-German *dorf*, likewise Low-German *dat* corresponds to High-German *das*, etc. However at the west end of the Bernarh-Urdingen line, this isogloss fans out into a number of intermediate isoglosses beginning at the Rhine (Rhenish fan). Isoglosses here separate the regions where a shift is at some intermediate stage, that is, it has affected some words but not yet others. In a region adjacent to a High-German region, people might say *apfel* for *appel*, but not *pfund* for *pund*. As one moves closer to the Low German region, fewer and fewer words are affected by the shift [8].

Changes spreading from a center are further thought to decay inversely with the distance. Lexical change, which appears to spread like this and also interact with boundaries such as mountains and oceans, implies that some innovations may never affect a distant group of speakers. Cultural isolation has in fact been observed to be a cause of language variety in many groups. The theory of *Isolation by Distance* or sometimes known as the *Stepping-Stone Model*, both aim to describe and understand how linguistic similarity varies with distance. There is a striking amount linguistic diversity amongst the many geographically separated islands that compose Southeast Asia and Oceania, exhibiting over 1400 of the 7000 known world languages. Data on certain Micronesian islands was collected to test this hypothesis. A correlation between distance and the number of shared cognates (words with similar roots) was observed to roughly exhibit a predicted exponential behavior [7].

B. The Micro Models

Thus far, the theories proposed to describe linguistic propagation have all been macroscopic in form. That is, they have aimed to understand diffusion of innovations in terms of the global observable quantities, such as the density of cognate similarities with distance, or isogloss diameter. While these macroscopic models are very useful to compare with real data, perhaps a model can be constructed in which macro-behavior emerges from known microscopic interactions between speakers, hereby recognizing the different scales from which the process may be viewed. From this consideration arose the *Neogrammarian Hypothesis* [14]. The idea was to describe the mechanism of diffusion on a smaller scale, in this case, on the level of a group of words connected to each other in some manner, such as a shared phoneme. According to the Neogrammarian Hypothesis, though changes in sound are phonetically gradual (the sound is successively pronounced with small alterations, the set of which composes a continuum of sound variants), the change is lexically abrupt (variation in pronunciation is not only seen in a single word, but in all words that share the change in question). This last condition implies that all words pertaining to a given group change simultaneously, though gradually with the differential sound shifts. A word group containing a particular phoneme thus undergoes a series of phonetic shifts before all words with that phoneme are pronounced with the different target phoneme.

This hypothesis neglects certain inhomogenous word mutations that have occurred in many languages. One such example in the English language, concerns the words *divine* and *divinity*. These two words once shared a common vowel sound in the location of the second syllable. However as language evolves, innovations are sometimes correlated with the presence, or lack thereof, of a given suffix or prefix. Thus, words which originally shared a common vowel sound, did not necessary evolve to have the same pronunciation [4]. In response, a second theory for the microscopic process of diffusion has been proposed. With *Lexical Diffusion* the primary difference is that word changes are not restricted to an entire word group but may act on single words. A lexically gradual change describes sound changes which occur word by word, from speaker to speaker, allowing words to change individually and not as an entire group as with lexical abruptness [14]. Further, phonetic changes may be gradual or abrupt with Lexical Diffusion. Since speakers tend towards a self-consistency, word pronunciations don't diverge in all possible directions. Due to this, one might understand how both theories are quite useful; recognizing that sound change shouldn't be restricted to act uniformly on an entire word group, but that there does tend to be change correlations among words of the same group.

As illustrated in [14], a sound change will invade the common lexicon at different rates, and not uniformly over a single word group. Word by word, and speaker by speaker, a sound change enters (and sometimes immediately exits) a language. Initially, few words participate in a sound change, but as more and more speakers use the change it becomes more linguistically accessible, and other similar words, in turn, undergo the sound change. The frequency of use for a given sound change increases within the quantity of words that are subject to the change, and thus invade the lexicon of speakers at an increasing frequency. This amplifying effect illustrates a phenomenon known as *The Snowball Effect* [6, 14]. To capture this process, a model presented in [6] begins by separating all changing word groups, each into two subpopulations: words that have undergone the change, and those that have not yet changed. For each word group, there is a population of speakers that use the unchanged form $u_i(t)$, and a population of speakers that use the changed form $c_i(t)$. At every differential time step, speakers choose to either adopt or not, a word change according to a particular dynamic. Should they change, the unchanged population decreases, and the changed population naturally increases. Both populations are normalized to one, $c_i(t) + u_i(t) = 1$. The chosen dynamic is that the unchanged population decreases at a rate proportional to the interaction of the existing unchanged population with the changed populations of all word groups within interaction probabilities β_{ij}, a coupling matrix.

$$\frac{du_i}{dt} = -u_i(t) \sum_{j=1}^{m} \beta_{ij}(1 - u_j(t))$$

In this model, all word groups m, of each speaker i, may interact with each other, but to varying degrees, designated by a coupling matrix [14]. One might surmise that the elements along the diagonal have the most weight, indicating that changed words are more likely to interact with unchanged words of the same word group. Given that the sum of all pairs of changed and unchanged populations for each word group is conserved, the dynamic is written as a logistic equation with a solution well-studied within physics and biology [5].

As mentioned previously, this work will focus on changes in vowel sounds for a population of speakers which share the same basic lexicon. We wish to explore the conditions necessary for accent diversification between speakers, as changes, innovations, and linguistic mutations propagate and diffuse through the population, speaker by speaker. Each speaker uses a set of word groups, each of which is characterized by a particular vowel sound. Each word group interacts with those of other speakers through a dynamic variable yet to be defined. Because of this, the process appears to be more like the Neogrammarian hypothesis. However, the concept might be extended to include lexically gradual changes as proposed in Lexical Diffusion. This currently is not within the scope of the project, however it ought not limit the interpretation of the following model and results.

IV. THE MODEL

A. Interaction Network

In a series of figures depicting regional differences in pronunciation of common English words across the United States, Labov illustrates the correlation between spatial location and accent diversity, see Fig. 1. One might ask, where does this variation come from? In this age of information and rapid communication, could such diversity be simply due to the geographic constraints that were imposed upon our ancestors, or is it a result of the social networks that humans weave amongst friends, collaborators, family, and media? Perhaps both, perhaps something else altogether. The roots of regional accent variation are likely to have begun before the pony express became the lean, mean, modern communication machine that it is today. Regardless, linguistic evolution continues, and language remains to be the transforming, mutating, propagating medium with which we express ideas in ways that are new to every generation.

In searching for explanations to these questions, we believe it is important to recognize the role played by the network of speaker interactions. Models often assume homogenous and well-mixed populations, such that interactions occur between any and all pairs of speakers. However, this mean field approximation is likely to wash out much of the behavior due to natural social phenomena. Speakers will here be considered to interact with other speakers within their neighborhoods, to varying degrees. A neighborhood is not a spatial restriction, but an abstraction of the members that a given speaker interacts with. One might further hope to consider interactions on a dynamic network, as relationships come and go, and as individuals are born and die. Though interesting, this dynamic nature complicates the system immensely. We propose that for a given time period, these variations in the population may be summarized within the respective weights between speakers [15].

The topology of the social network is thus an important feature to explore in the process of linguistic diversification. We have chosen to use a Small World Network (SWN) to represent the underlying social structure of speakers [16]. However, a more extensive analysis in terms of the order parameter p (the level of disorder in the network) ought to be examined in further detail and most of the numerical analysis has been done on a random $p = 1$ graph. Each speaker occupies a node on the network. Edges between nodes indicate that speakers will interact socially, that is, through language. Since all speakers share the same lexicon, words may be divided into word groups (or syllable groups), that share the same vowel sound. The most obvious form of regional accent is in different pronunciations of common vowel sounds. We hope that such a division is sufficient to capture the most marked differences in accent. In general, all word groups of connected speakers may interact with all other word groups (as in the multidimensional version of the Snowball Effect [14]). In the model presented here, we again try to extract the most important interactions (specifically, maximization of speaker-listener efficiency), and restrict interactions to be between similar word groups.

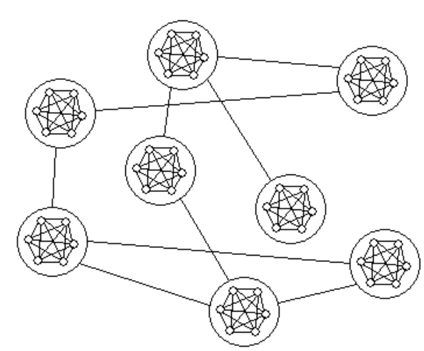

FIG. 2: Larger circles represent speakers, while smaller word groups. Within every speaker there is an interconnected network of word groups that interact with each other. As one word group changes, it will change others (e.g. to prevent merging or word groups). Thus, there are two kinds of interactions, between speakers, and within a speaker.

Among the many factors that influence language, linguists argue that two prominent ‘forces’ are born from the need to efficiently communicate. In an attempt to reduce energy and effort, speakers tend to simplify language. Whereas, listeners will demand clarification and annunciation, if the message is not understood. A balance between these two effects is thus sought [4]. We will draw more attention to this attribute later, however we now summon the concept to motivate what occurs in a speaker’s own vocabulary. While specific word groups among speakers are here only allowed to interact with like-word groups of other speakers, for a single speaker, all of the different word groups will indeed interact amongst each other. The intention is to recognize the self-consistency that a speaker maintains. Should one’s word groups only be affected by other speakers, and not at all by one’s own manner of speaking, then the word groups contained by this individual might merge to form one indistinguishable mess of the same sound. Thus, we find it natural to require that a speaker is wary of how he/she pronounces the other words within his/her lexicon. Should vowel sound changes occur, the speaker will make an effort to be understood by distinguishing word groups, however slightly. A piece of the speaker-word group network described, is illustrated in Fig. 2.

B. Interaction Dynamic

Between the centuries 1200 and 1800, the English language was undergoing a profound change [2]. Vowel phonemes in nearly all words were drifting to a different pronunciation. Though the shift took about 600 years to complete, the differences between modern English and Ye Olde Englishe are deep. To explain exactly how this Great Vowel Shift took place, we begin with a diagram of the mouth, Fig. 3. Each vowel sound, and each sound for that matter, is produced by a particular combination of air release, vocal chord contraction, mouth configuration, and tongue placement. Through observation and practice, linguists have mapped out the necessary muscular formations to produce such sounds. Vowel phonemes are different for every language, yet each have their place within the mouth. During the great vowel shift, vowels shifted primarily along two inward spiraling circles in the front of the mouth and in the back of the mouth. It is interesting to note that changes occurred cyclically and phonetically gradually. Recall that for the Neogrammarian Hypothesis, as well as some cases of Lexical Diffusion, changes occur in a phonetically gradual manner. In relation to the mouth, this means that a vowel sound moves by small differential steps from one location in the mouth to another. This is a key concept in the dynamic that we now propose.

Given the structure of the mouth, is it possible to extract a variable that might dynamically track the pronunciation of vowel sounds pertaining to a given word group? If we want to measure the interaction between two sounds that are very close, we might characterize

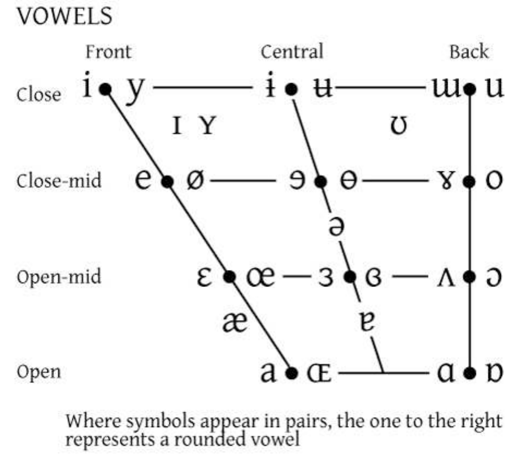

FIG. 3: The vowel sounds of English are produced in different parts of the mouth as depicted here. The left side of the figure represent the front of the mouth. This figure may be reproduced with acknowledgments to the International Phonetic Association (Department of Theoretical and Applied Linguistics), School of English, Aristotle University of Thessaloniki.

the vowel sounds with a linear variable. That is, if differences between the two vowel sounds under examination are small, such that by zooming into a region of the mouth the variation between two sounds is not periodic, then the sounds may be mapped onto a line. Further, they should approach each other monotonically; that the changing pronunciation doesn’t drastically overshoot that target pronunciation. A linear mapping may work fine for small shifts, but what if we want to map the interactions and resulting shifts between all vowel sounds in all word groups? In this case, we cannot simply zoom into a small region and map to the line, but must consider a mapping realistic to the proper scale. Here, we suggest a mapping of the mouth onto the unit circle. This is a useful approximation because every vowel sound may now be characterized by a unique angle within the 2π spectrum, and is called a phase. Through a judiciously chosen interaction function, phase variables can now interact and influence each other, and thus return a possible set of pronunciations for different word groups, given social interactions and certain coupling conditions yet to be discussed.

To proceed with the analysis, we must first state some of the restrictions that are imposed by characterizing the pronunciation with a phase variable. First of all, not all vowel sound changes occur in such a continuous manner within the mouth. Specifically, those changes which are considered to be lexically abrupt. Secondly, we assume that the speaker can only transition from one vowel sound to a single other vowel sound. This cyclic requirement was motivated by the cyclic transitions observed in the Great Vowel Shift [2]. Accessibility between vowel sounds is thus analogous to the accessibility between points on a circle. One revolution around the circle would thus correspond to cycling through the entire vowel sound set. In terms of a phase variable, any transition adding or subtracting an integer multiple of 2π radians will leave a vowel sound unchanged. In summary, we assume that changes in vowel sound pronunciation are both *cyclic* and *continuous*.

C. Coupling the Interactions

Now that the two levels, the social interaction network and the internal dynamic phase variables, have been introduced, we may now begin to put these two concepts together as a working model. Every speaker is connected to a small (in relation to total network size) group of other speakers. These speakers interact with each other through distinct word groups, each characterized by a single phase variable. As an example to illustrate this, consider a particular set of sounds present in the canonical Boston accent. In this accent, words such as 'car' or 'bar' sound, respectively, like 'cah' and 'bah'. (While the illustrative example here concerns a sound that is typically characterized by consonant pronunciation of the r, in the model will focus solely on vowel sound pronunciations). As the Bostonian interacts with people from other regions of the country, how will the accent be affected? We first assume that the speaker is relatively self-consistent, that is, it is unlikely that he/she says 'cah' and 'bah' with a perfect Boston accent, while otherwise maintaining a flawless Southern drawl. Individuals strive, both consciously and unconsciously, to maintain self-consistency. Innovative vowel sounds will thus be checked with other word group pronunciations within their own vocabulary. With that said, the method of vowel sound diffusion is through the conscious and unconscious experimentation of a new pronunciation. If a Bostonian begins to imitate a sound present on the west coast, such as within the 'car/bar' word group, the innovation will enter through the single 'cah/bah' word group (or through a single word, which again isn't distinguished in the current model). It will then be accepted (the change remains and may affect other internal word groups such as 'cord/sword'), or declined (change is washed out) by the other word groups within his/her lexicon, such as those which are characterized by the sounds 'caught/daugther', or 'cord/sword', among many others. Through the phase variable of each of these word groups, speakers influence and interact with each other, accepting and declining innovations.

A further point to consider within the interactions among speakers are the social pressures discussed in previous sections. In the model, we consider various coupling matrices and a single interaction function. We do this in a general manner for the present work, and neglect to analyze how certain social constraints (such as, prestige, gender, age, etc.) may motivate these two parameters (interaction function and coupling matrix), or vary spatially within the network topology. Here we have assumed that speakers all have equal prestige, where there are no social classes that differentiate one from another. Gender nor age specifically influence innovation propagation, and finally, there is no standardized language nor centralized reinforcement of a particular linguistic form. To some extent, these may be encoded in the coupling matrix, however the bias towards higher social class and prestige is entirely neglected for the time being. By adding pertinent features to the present idealized model, we hope to later address how social pressures change the overall dynamics and in particular, the equilibrium state, should it exist. Presently, we have only imposed the condition that speakers and listeners wish to communicate efficiently.

D. The Mathematical Formulation

Though we have already outlined our assumptions and restrictions on the model, we here begin by writing the most general, mathematical interaction representation. This may be useful for future work, where different interactions (non pair-wise, between particular word groups, different interaction functions, etc.) might want to be considered. The dynamic variable is initially written in vector form, each component pertaining to a different word group, i.e. $\vec{r}_i = (\varphi_i^1, \varphi_i^2, ..., \varphi_i^m)$. We thus consider how the pronunciation of a given speaker i, is changing through the dynamic interactions with all other speakers 1 thru n.

$$\dot{\vec{r}}_i = \vec{F}_i(\vec{r}_i, \vec{r}_1, ..., \vec{r}_n)$$

This composes the set of differential equations for all m word groups of a single speaker. Taking into account all n speakers, we find that in general, the system is composed of $n \times m$ equations, each with a potentially different interaction function (this is why F has also been written in vector form for each speaker). This beastly mess of equations has been expressed here simply to suggest the degree of detail that one may elect, and further, to point out where our assumptions and restrictions begin to reduce the most general form. To reduce this system of equations to something manageable, lets begin by considering the dynamic to be composed of two types of interactions, 1) *internal* couplings: the couplings between the word groups of the speaker in question, $\vec{r}_i$, and 2) *external* couplings: the couplings with all other speakers, $\vec{r}_1$ thru $\vec{r}_n$. For certain restrictions outlined in [17–19], this system may be examined analytically with a very general and elegant stability analysis method. Linguistically, this potentially corresponds to studying how susceptible a given pronunciation is to invading accents, however, we reserve this analysis for a future work.

We now restrict the interactions to be only pair-wise. This means that the phase variable φ_i, of a single word group of a single speaker, only interacts with a single other phase variable. The dynamic of this phase variable $\dot{\varphi}_i$ is proportional to the sum of all such pairwise interactions. Further assuming that all phase variables have the same interaction function, the dynamic of a single word group of speaker i is immediately reduced to the form,

$$\dot{\varphi}_i = \sum_{j=1}^{n}\sum_{k=1}^{m} F(\varphi_j^k, \varphi_i)$$

where the first sum is over speakers, and the second sum is over word groups. This coupling scheme is a general form for any pair-wise interacting variables. For periodically varying dynamic variables, as is our case since we have mapped the anatomical vowel sound chart onto the circle, there is a well studied formulation originally given in [20], and explained briefly in layman's terms in [5]. The final interaction equation is derived such that the rate of phase change is proportional to the difference of phase variables.

$$\dot{\varphi}_i = \frac{1}{N}\sum_{j=1}^{n}\sum_{k=1}^{m} A_{ij}^m sin(\varphi_j^k - \varphi_i)$$

The multiplicative constant A_{ij}, is the coupling strength between two phase variables. Variations in this value imply that certain word groups affect some word groups more than others. To illustrate this, possible coupling constants for a three speaker, two word group system has been written below,

$$\mathbf{A} \rightarrow \left(\begin{array}{ccc} \begin{pmatrix} \alpha & \alpha \\ \alpha & \alpha \end{pmatrix} & \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix} & \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix} \\ \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix} & \begin{pmatrix} \alpha & \alpha \\ \alpha & \alpha \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} \alpha & \alpha \\ \alpha & \alpha \end{pmatrix} \end{array} \right)$$

The matrices along the diagonal represent the internal couplings of a speaker. Each α corresponds to a coupling strength between word groups. The off diagonal matrices represent couplings between different speakers. The β values are only along the diagonals to indicate the restriction that we have imposed, only like word groups among different speakers will affect each other. That is, two speakers may interact through 'car/bar' and 'cah/bah', but they will not interact through 'cord/sword' and 'cah/bah'. The matrices composed entirely of zeros simply indicate that speakers 2 and 3 are not connected. Finally, note that the entire matrix is symmetric, indicating the assumption that speakers affect each other to the same degree. The coupling values α and β are in general, random variables chosen from a given distribution, and for now maintain the symmetry. We will discuss possible distributions and the subsequent random values, in the numerical analysis section.

As pronunciations become more alike, the difference of phase variables goes to zero. One is imitating the other with finer accuracy. However, communication is hardly ever perfect. To capture these misinterpretations and misutterances, we introduce an error into what is heard. The speaker i hears a pronunciation of the k^{th} word group of the j^{th} speaker as $\varphi_j^k \longrightarrow \varphi_j^k + \epsilon_{ij}^k$, for reasons of misinterpretation, anatomical differences, or a lack of familiarity with the accent, etc. If these errors are considered to be linear additions, then the difference of phase variables no longer converges on zero, but on the error. For a single word group of speaker i, the final model to be analyzed is thus,

$$\dot{\varphi}_i = \frac{1}{N}\sum_{j=1}^{n}\sum_{k=1}^{m} A_{ij}^m sin(\varphi_j^k - \varphi_i - \epsilon_{ij}^k) \qquad (1)$$

E. Pronunciation Synchronization

The evolutionary precursors that led to linguistic communication are still debated among linguistics, though few would argue that some anatomical changes were necessary. Also, somewhere within our history, humans acquired the interest in naming things [21]. Children show an immense adeptness for this tendency; names and their associations to meanings are imitated from parents and other peers. These actual meaning utterance maps, of course vary from nation to nation, though the capability remains universal to all modern human beings. Language has played a leading role in the development of political and cultural boundaries, literary and technological styles, cultural perceptions, etc. These examples may be seen as bits of strong macroscopic evidence that a special sort of organization has occurred in the human species, and continues to occur as younger generations are born into this communicative organism. The organized and diverse human languages that we see today, have been based simply on the microscopic interactions between the original speakers, between children and their parents, between those learning a new language through conversation. Within many complex systems, this behavior loosely indicates that a sort of self-organization has occurred [5, 16, 21]. A globally organized state which emerges due solely to local interactions. It may thus be feasible to consider language as a synchronization phenomena.

In the synchronization of regional accent, we ask not only whether speakers will adopt each other's pronunciation, but if the population maintains a diversification of sounds. Accent diversity naturally ought to follow from many social particularities: the migration of people speaking foreign languages; the influence of the media; the inertia or proclivity to change in different personalities; the mixing of generations and their relative adaptability; and the topology of social interactions. The model as such currently addresses a subset of the above. We have chosen to look at the combined effects of social topology, interaction strength between speakers, and errors in reception and reproduction of sound. In function of these three parameters, we hope to measure how global pronunciation schemes form and persist, given only the local interactions between speakers. An order parameter, discussed in [5, 20] and elsewhere, is useful to characterize the organization within the system. It measures how synchronized is a group of speakers, as they pronounce a given word group. Correspondingly, the parameter also depicts the diversity present among speaker pronunciations. This quantity is defined by the magnitude $R(t)$, of the averaged vector sum of all phase variables within a single word group,

$$R(t)e^{i\Phi(t)} = \frac{1}{n}\sum_{j=1}^{n} e^{i\varphi(t)}$$

That is, each point on the circle in Fig. 4, is characterized by a complex vector, with polar coordinates given by the unit radius and the phase variable. When all pronunciations (phase variables) are dispersed evenly throughout the unit circle, the vectors average to zero. As speakers interact, the phase variables become more similar, and pronunciation clusters form, corresponding to the different word groups. As this happens, the phases of all complex vectors cease to cancel, and the magnitude grows toward unity.

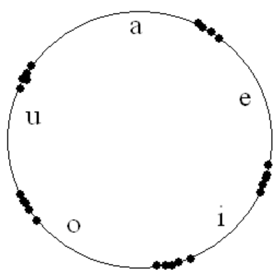

FIG. 4: Each point represents the vectorial position of each phase variable on the unit circle mouth mapping. As pronunciations synchronize, phase variables of the same word group tend to cluster around the accepted pronunciation. A spread within the cluster indicates a certain level of diversity in the pronunciation of word groups.

V. NUMERICAL ANALYSIS

The dynamics of the system contained in eq.(1) have been numerically solved from $t = 0$ until the solution stabilizes. Pertinent parameters have be set to be,

- There are $n = 100$ speakers, each communicating with $m = 7$ word groups.
- The network of speakers is a random network with degree $k = 10$, thus on average each speaker has 10 neighbors. The network is non-directed implying that the adjacency matrix is symmetric.
- The internal coupling α matrices are set to be stronger than the external coupling β matrices. Specific values for the α matrices are considered in the following sections. The constraint that speakers only interact with the same word groups of their neighbors implies that the β matrix is a diagonal, where diagonal terms decreases exponentially according to Zipf's law.
- Imitation error ϵ_{ij}^{k} is chosen from a normal distribution scaled down so that it's variance is twice as small as the smallest external coupling.
- The initial phases are picked randomly, i.e. uniform over the interval $(-\pi, \pi)$.
- A fourth order Runge-Kutta method has been used to integrate the differential equation, which has been shown to exhibit instabilities for lower order methods.

In the present model we have not explicitly analyzed the effects of different network topologies. Some exploratory investigations indicate that the pronunciations more readily synchronize with more disorder, although we are not sure how the disorder interplays with different coupling matrices. There is a complete analysis of synchronization phenomena in disordered and ordered networks in [22]. We plot both the phase variable evolution and the synchronization order parameter R(t) for several novel cases. Through different coupling configurations that pertain to possible social situations, we explore the nature of the system defined by our model. The coupling matrix is the primary tool with which we can do this. In all cases examined below, we choose the same couplings between word groups of neighboring speakers. The motivation behind these β values, lies within Zipf's law. It has been stated that certain word groups are used with a frequency that diminishes as a power law. We begin with a trivial case to illustrate the effect of Zipf's law, and further properties presented in the model. Another important issue is the relative strength of the internal alpha couplings to the external beta couplings. Given our lack of linguistic background, we conjecture that internal couplings are stronger than the couplings among speakers.

A. A Trivial Case

In Fig. 5, though phase variables begin with a random distribution, they all approach the same exact value. This implies the startling condition that all words are pronounced with the same vowel sound. In this simulation, we have allowed all α and β couplings to be positive. We pick α to be a constant matrix with each entry twice larger then the largest value in β. Given that there is no natural frequency, nor linguistically analogous term for the time being, all phase variables approach each other as one might expect [5, 22]. This is also observed in the synchronization plot Fig. 6, where phase variables for each word group approach a synchronized pronunciation. The dispersal of the curves is due to the different values of β demonstrating the effect that Zipf's law has on the synchronization process. Note that regardless of the network topology the same equilibrium state is acquired. Although the transient effects may linger in weakly connected networks, in the infinite time limit the steady state is the same.

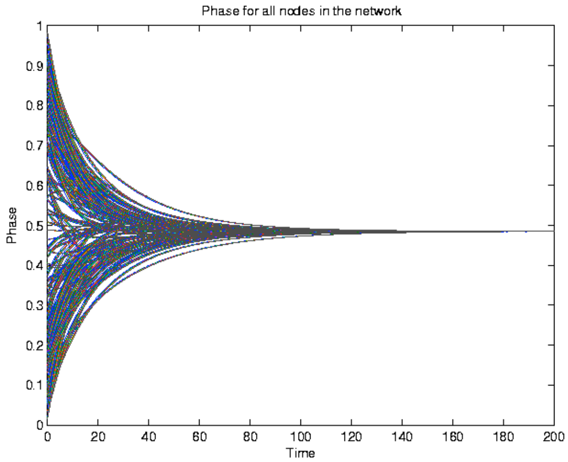

FIG. 5: The evolution of all phase variables in the system converge in time. Each word group is the same color for all speakers. When both α and β matrices are contain positive elements, we achieve a global synchronization, thus all word groups are pronounced the same in long enough time.

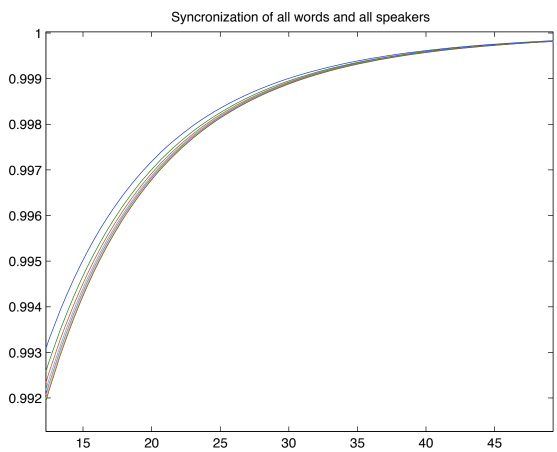

FIG. 6: Every curve depicts a different word group. Due to the Zipf's law, we see a dispersion among word groups as each progresses towards synchronization. Both α and β matrices have positive components.

B. Negative Internal Couplings

We now take into consideration one of the linguistic necessities that has been mentioned several times already. This is the need to be understood, to enunciate, to distinguish words for the sake of the listeners. Due to this, we have required that word groups are distinct from each other. Though mergers (words that acquire the same pronunciation) occur especially in the English language, we neglect this possibility for the moment because it tends to occur for single words. The necessity of word group distinction thus implies a sort of repulsive force between word groups of a given speaker. For this reason, we now take the symmetric internal α couplings to be negative numbers, each chosen from a uniform distribution. The external couplings are still positive and chosen from a power law distribution. We find that while word groups are now differentiated, they still converge onto a single pronunciation, see Fig. 7. All speakers now have a varied vocabulary, but they speak with exactly the same accent. No vowel sound diversity arises in this case. This is also observed in Fig. 8, where all word groups synchronize completely. Again, regardless of the network topology, we see pronunciation synchronization, however, in this case synchrony is observed among words groups, and not between all words in the lexicon.

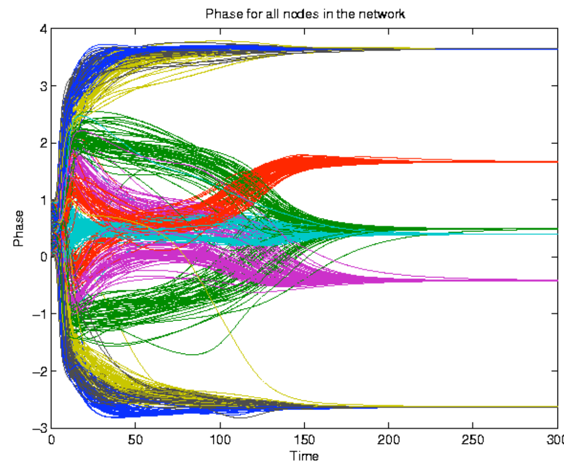

FIG. 7: The evolution of all phase variables in the system converge in time for a negative coupling α matrix. Each color depict a different word groups. We see that word groups separate from other groups and synchronize within their group. Keep in mind that phases vary with a period of 2π, thus the lower blue word group is the same as the upper blue word group.

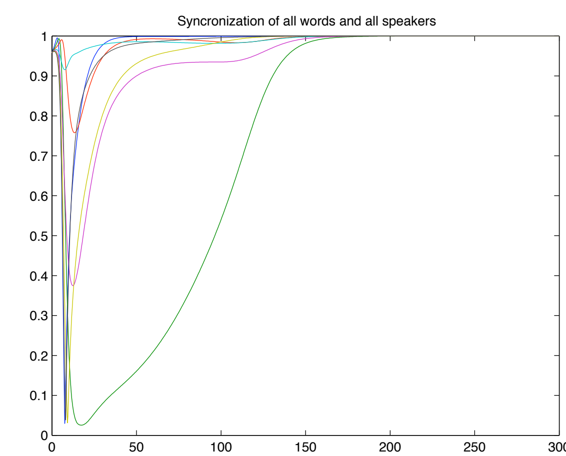

FIG. 8: Synchronization for a negative coupling α matrix. After a transient period we reach synchronization within each word group.

C. Distinct Individual Internal Couplings

Distinctions between individual speakers may translate into different relative weights between word groups, as speakers demonstrate different preferred vocabularies, different interactions between pronunciations, word groups which may be heavily coupled through slang or regional culture, etc. Given the unique conditions of all speakers, we now chose the internal interaction matrices to be symmetric and negative as before, but now each α matrix is different for every speaker. In this case, we recover some interesting behavior, see Fig. 9. Word groups tend to cluster and synchronize, but the speakers don't all converge onto the exact same sound. That is, pronunciation groups form, but there is a spread in the vowel sound amongst the speakers, and we thus see diversity of accent. Again, the synchronization profiles of Fig. 10 show how word groups approach each other, but do not fully synchronize. Some word groups show much more diversity of vowel sound pronunciation than other word groups, as one might expect to see in real linguistic systems.

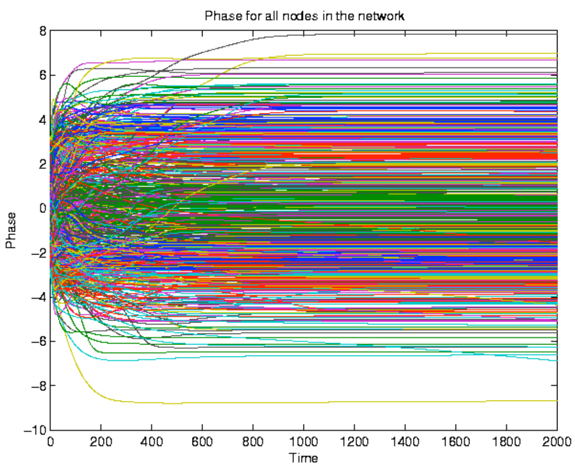

FIG. 9: The evolution of all phases in the system when α has both positive and negative elements, but is different for every speaker. Observe the pronunciation diversification driven by individual preferences and the necessity to communicate (i.e. synchronize with neighbors).

D. Same Internal Couplings with Imitation Error

We now take into account that every speaker-listener pair is assigned an imitation (reproduction) error for each word group, as expressed in eq.(1). Since we consider only pronunciation error of speakers, we set that every speakers has only one pronunciation error regardless of the listener. Thus ϵ_{ij}^{m} is same for all i, and ϵ_{ii} is zero vector (since internal coupling are not influenced by the imitation error). The imittation error values are chosen from a normal distribution with zero mean value, and

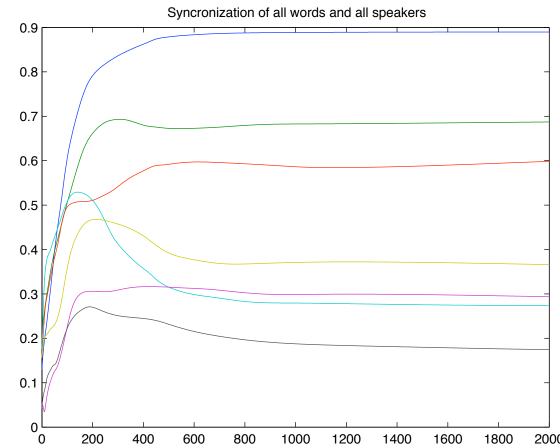

FIG. 10: Synchronization for the α coupling matrix is different for every speaker. Word groups synchronize, each to a different extent.

a variance twice smaller than the external smallest couplings. We set the α matrices to be negative as done previously, to give rise to distinct word groups, yet the same for every speaker, see Fig. 11. Even in a random network, imitation error induces a spread around the synchronized state. It is left for future work to see if with highly clustered (and often disconnected) network architecture, one would sees an accumulation of imitation error, and thus an even larger spread around the synchronized state, as the mean path through the network increases (with disconnection).

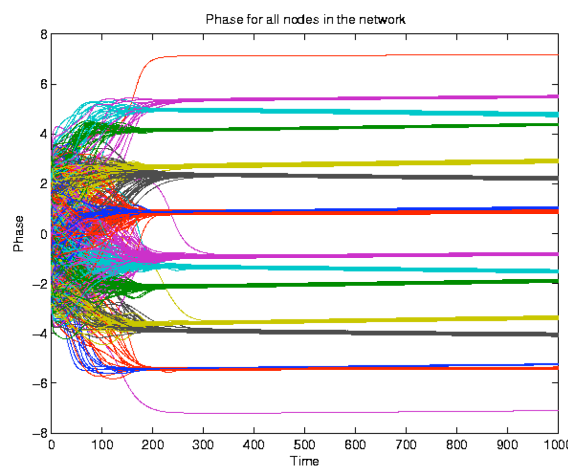

FIG. 11: Phase evolution in the system in the presence of imitation error. Each color depicts a different word group. Here, the spread around the synchronized state is due to imitation error. Recall that phases vary within a period of 2π, thus lower and upper yellow are same phase.

VI. CONCLUSIONS

The self-organized behavior among human beings, known as language, is a highly complex phenomena that perhaps may only be modelled in steps, as the assumptions and restraints may change wildly among each linguistic property one elects to study. Here we have attempted to model the diversification of regional dialect through the social interactions of vowel sound pronunciation of speakers confined to a single lexicon. Before exploring this idea, a small portion of the multitudes of similar linguistic work was sampled and presented here as the preceding theories that have pioneered linguistic diffusion and propagation. With this background, we have built a model for linguistic interaction based primarily on two principals: 1) that speakers interact through a network of social connections through a vocabulary divided into word groups. The manner in which these word groups interact is through imitation of a phase variable which is, 2) extracted from the mapping of the vowel sound mouth chart, onto the unit circle. For a very simplified interaction function (which may change to incorporated prestige, age, and gender pressures among others), we have iterated the proposed linguistic dynamic on the network, for various coupling matrices and imitation errors. Some of the couplings produce no novel behavior, whereas others lead to an apparent diversification of language, that maintains some synchronization so that speakers may still communicate. Clusters of similar pronunciations for different word groups are found to form when 1) the internal individual coupling matrices are all different, symmetric, and negative, and 2) when there is a nonzero reception/imitation error between speakers. More analysis and discussion is necessary to determine if this is interesting to actual linguistic research or just an artifact of the model. We suspect that network topology and certain social pressure ought to play a formative role in accent diversification as well. Finally, it may be interesting to conduct an stability analysis of these accent clusters to see how the population of speakers responds to pronunciation perturbations on a large scale, perhaps associated with migration or exodus.

VII. ACKNOWLEDGEMENTS

The authors would like to thank the Santa Fe Institute and the Chinese Academy of Science for hosting the Complex Systems Summer School 2005, where this work was conducted. We would also like to thank William S.Y. Wang for helpful discussion and encouragement.

[1] J. Plotkin, M. Nowak. Entropy **3**, 227-246 (2001).
[2] http://www.courses.fas.harvard.edu/ laa14/vowels.html
[3] M. Ruhlen. *The Origin of Language*, John Wiley & Sons (1994).
[4] conversations with W.S-Y. Wang, Complex Systems Summer School, hosted by SFI and CAS-ITP (2005).
[5] S. Strogatz *Nonlinear Dynamics*
[6] W.S-Y. Wang, J. Ke., J. Minett
[7] L. Cavalli-Sforza, W.S-Y. Wang. Language **62** 1, 38-55 (1986).
[8] S. Romaine *Language in Society* 2nd ed. Oxford Press, (2000).
[9] W. Labov, linguistic reference website: http://www.ling.upenn.edu/ wlabov/home.html
[10] W.S-Y. Wang, J. Minett. Transactions of the Philological Society, **103**, 121-146 (2005).
[11] K. Smith, S. Kirby, H. Brighton. Artificial Life **9**, 371-386 (2003).
[12] N. Eliasoph, P. Lichterman. Amer. Journ. of Sociology **108** 4, 735-94 (2003).
[13] Nationmaster, online encyclopedia: http://www.nationmaster.com/encyclopedia/Code_switching
[14] W.S-Y. Wang, J. Minett. TRENDS in Ecology and Evolution **20**, 264-269 (2005).
[15] D. Stillwell, E. Bollt, D. Roberson. arxiv.org/abs/nlin.CD/0502055
[16] D. Watts. *Small Worlds, The Dynamics of Networks between Order and Randomness* 223-239. Princeton University Press (1999).
[17] K. Fink, G. Johnson, T. Carroll, D. Mar, L. Pecora. Phys. Rev. E **61**, 5080-5090 (2000)
[18] L. Pecora. Phys. Rev. E **58** 347-360 (1998).
[19] L. Pecora, T. Carroll. Phys. Rev. Lett. **80** (1998).
[20] Y. Kuramoto, *Proceedings of the International Symposium on Mathematical Problems in Theoretical Physics*. Springer, New York (1975).
[21] J. Ke, J. Minett, C.P. Au, W.S-Y. Wang. Complexity (2002).
[22] S. Manrubia, A. Mikhailov, D. Zanette. *Emergence of Dynamical Order*, World Scientific (2004).